



PCT10

RAW SEQUENCE LISTING

DATE: 06/05/2002

PATENT APPLICATION: US/10/031,146

TIME: 17:09:35

Input Set : A:\031146.txt

Output Set: N:\CRF3\06052002\J031146.raw

Does Not Comply
Corrected Diskette Needed

3 <110> APPLICANT: HAUER, Bernhard
 4 PLEISS, Juergen
 5 SCHWANEBERG, Ulrich
 6 SCHMITT, Jutta
 8 <120> TITLE OF INVENTION: <120> Novel cytochrome P450 monooxygenases and their use for
 the oxidation of
 9 organic substrates
 11 <130> FILE REFERENCE: 50915
 13 <140> CURRENT APPLICATION NUMBER: 10/031,146
 C--> 14 <141> CURRENT FILING DATE: 2002-04-30
 16 <150> PRIOR APPLICATION NUMBER: PCT/EP 00/07253
 17 <151> PRIOR FILING DATE: 2000-07-27
 19 <160> NUMBER OF SEQ ID NOS: 9
 21 <170> SOFTWARE: Word Perfect version 6.1
 23 <210> SEQ ID NO: 1
 24 <211> LENGTH: 3150
 25 <212> TYPE: DNA
 26 <213> ORGANISM: Bacillus megaterium
 28 <220> FEATURE:
 29 <221> NAME/KEY: CDS
 30 <222> LOCATION: (4)..(3150)
 32 <400> SEQUENCE: 1
 34 atg aca att aaa gaa atg cct cag cca aaa acg ttt gga gag ctt aaa 48
 35 Thr Ile Lys Glu Met Pro Gln Pro Lys Thr Phe Gly Glu Leu Lys
 36 1 5 10 15
 38 aat tta ccg tta tta aac aca gat aaa ccg gtt caa gct ttg atg aaa 96
 39 Asn Leu Pro Leu Leu Asn Thr Asp Lys Pro Val Gln Ala Leu Met Lys
 40 20 25 30
 42 att gcg gat gaa tta gga gaa atc ttt aaa ttc gag gcg cct ggt cgt 144
 43 Ile Ala Asp Glu Leu Gly Glu Ile Phe Lys Phe Glu Ala Pro Gly Arg
 44 35 40 45
 46 gta acg cgc tac tta tca agt cag cgt cta att aaa gaa gca tgc gat 192
 47 Val Thr Arg Tyr Leu Ser Ser Gln Arg Leu Ile Lys Glu Ala Cys Asp
 48 50 55 60
 50 gaa tca cgc ttt gat aaa aac tta agt caa gcg ctt aaa ttt gta cgt 240
 51 Glu Ser Arg Phe Asp Lys Asn Leu Ser Gln Ala Leu Lys Phe Val Arg
 52 65 70 75
 54 gat ttt gca gga gac ggg tta ttt aca agc tgg acg cat gaa aaa aat 288
 55 Asp Phe Ala Gly Asp Gly Leu Phe Thr Ser Trp Thr His Glu Lys Asn
 56 80 85 90 95
 58 tgg aaa aaa gcg cat aat atc tta ctt cca agc ttc agt cag cag gca 336
 59 Trp Lys Lys Ala His Asn Ile Leu Leu Pro Ser Phe Ser Gln Gln Ala
 60 100 105 110
 62 atg aaa ggc tat cat gcg atg atg gtc gat atc gcc gtg cag ctt gtt 384

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/031,146

DATE: 06/05/2002

TIME: 17:09:35

Input Set : A:\031146.txt

Output Set: N:\CRF3\06052002\J031146.raw

```

63 Met Lys Gly Tyr His Ala Met Met Val Asp Ile Ala Val Gln Leu Val
64      115      120      125
66 caa aag tgg gag cgt cta aat gca gat gag cat att gaa gta ccg gaa 432
67 Gln Lys Trp Glu Arg Leu Asn Ala Asp Glu His Ile Glu Val Pro Glu
68      130      135      140
70 gac atg aca cgt tta acg ctt gat aca att ggt ctt tgc ggc ttt aac 480
71 Asp Met Thr Arg Leu Thr Leu Asp Thr Ile Gly Leu Cys Gly Phe Asn
72      145      150      155
74 tat cgc ttt aac agc ttt tac cga gat cag cct cat cca ttt att aca 528
75 Tyr Arg Phe Asn Ser Phe Tyr Arg Asp Gln Pro His Pro Phe Ile Thr
76 160      165      170      175
78 agt atg gtc cgt gca ctg gat gaa gca atg aac aag ctg cag cga gca 576
79 Ser Met Val Arg Ala Leu Asp Glu Ala Met Asn Lys Leu Gln Arg Ala
80      180      185      190
82 aat cca gac gac cca gct tat gat gaa aac aag cgc cag ttt caa gaa 624
83 Asn Pro Asp Asp Pro Ala Tyr Asp Glu Asn Lys Arg Gln Phe Gln Glu
84      195      200      205
86 gat atc aag gtg atg aac gac cta gta gat aaa att att gca gat cgc 672
87 Asp Ile Lys Val Met Asn Asp Leu Val Asp Lys Ile Ile Ala Asp Arg
88      210      215      220
90 aaa gca agc ggt gaa caa agc gat gat tta tta acg cat atg cta aac 720
91 Lys Ala Ser Gly Glu Gln Ser Asp Asp Leu Leu Thr His Met Leu Asn
92      225      230      235
94 gga aaa gat cca gaa acg ggt gag ccg ctt gat gac gag aac att cgc 768
95 Gly Lys Asp Pro Glu Thr Gly Glu Pro Leu Asp Asp Glu Asn Ile Arg
96 240      245      250      255
98 tat caa att att aca ttc tta att gcg gga cac gaa aca aca agt ggt 816
99 Tyr Gln Ile Ile Thr Phe Leu Ile Ala Gly His Glu Thr Thr Ser Gly
100      260      265      270
102 ctt tta tca ttt gcg ctg tat ttc tta gtg aaa aat cca cat gta tta 864
103 Leu Leu Ser Phe Ala Leu Tyr Phe Leu Val Lys Asn Pro His Val Leu
104      275      280      285
106 caa aaa gca gca gaa gaa gca gca cga gtt cta gta gat cct gtt cca 912
107 Gln Lys Ala Ala Glu Glu Ala Ala Arg Val Leu Val Asp Pro Val Pro
108      290      295      300
110 agc tac aaa caa gtc aaa cag ctt aaa tat gtc ggc atg gtc tta aac 960
111 Ser Tyr Lys Gln Val Lys Gln Leu Lys Tyr Val Gly Met Val Leu Asn
112      305      310      315
114 gaa gcg ctg cgc tta tgg cca act gct cct gcg ttt tcc cta tat gca 1008
115 Glu Ala Leu Arg Leu Trp Pro Thr Ala Pro Ala Phe Ser Leu Tyr Ala
116 320      325      330      335
118 aaa gaa gat acg gtg ctt gga gga gaa tat cct tta gaa aaa ggc gac 1056
119 Lys Glu Asp Thr Val Leu Gly Gly Glu Tyr Pro Leu Glu Lys Gly Asp
120      340      345      350
122 gaa cta atg gtt ctg att cct cag ctt cac cgt gat aaa aca att tgg 1104
123 Glu Leu Met Val Leu Ile Pro Gln Leu His Arg Asp Lys Thr Ile Trp
124      355      360      365
126 gga gac gat gtg gaa gag ttc cgt cca gag cgt ttt gaa aat cca agt 1152
127 Gly Asp Asp Val Glu Glu Phe Arg Pro Glu Arg Phe Glu Asn Pro Ser

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/031,146

DATE: 06/05/2002

TIME: 17:09:35

Input Set : A:\031146.txt

Output Set: N:\CRF3\06052002\J031146.raw

128	370	375	380	
130	gcg att ccg cag cat gcg ttt aaa ccg ttt gga aac ggt cag cgt gcg	1200		
131	Ala Ile Pro Gln His Ala Phe Lys Pro Phe Gly Asn Gly Gln Arg Ala			
132	385 390 395			
134	tgt atc ggt cag cag ttc gct ctt cat gaa gca acg ctg gta ctt ggt	1248		
135	Cys Ile Gly Gln Gln Phe Ala Leu His Glu Ala Thr Leu Val Leu Gly			
136	400 405 410 415			
138	atg atg cta aaa cac ttt gac ttt gaa gat cat aca aac tac gag ctg	1296		
139	Met Met Leu Lys His Phe Asp Phe Glu Asp His Thr Asn Tyr Glu Leu			
140	420 425 430			
142	gat att aaa gaa act tta acg tta aaa cct gaa ggc ttt gtg gta aaa	1344		
143	Asp Ile Lys Glu Thr Leu Thr Leu Lys Pro Glu Gly Phe Val Val Lys			
144	435 440 445			
146	gca aaa tcg aaa aaa att ccg ctt ggc ggt att cct tca cct agc act	1392		
147	Ala Lys Ser Lys Lys Ile Pro Leu Gly Gly Ile Pro Ser Pro Ser Thr			
148	450 455 460			
150	gaa cag tct gct aaa aaa gta cgc aaa aag gca gaa aac gct cat aat	1440		
151	Glu Gln Ser Ala Lys Lys Val Arg Lys Lys Ala Glu Asn Ala His Asn			
152	465 470 475			
154	acg ccg ctg ctt gtg cta tac ggt tca aat atg gga aca gct gaa gga	1488		
155	Thr Pro Leu Leu Val Leu Tyr Gly Ser Asn Met Gly Thr Ala Glu Gly			
156	480 485 490 495			
158	acg gcg cgt gat tta gca gat att gca atg agc aaa gga ttt gca ccg	1536		
159	Thr Ala Arg Asp Leu Ala Asp Ile Ala Met Ser Lys Gly Phe Ala Pro			
160	500 505 510			
162	cag gtc gca acg ctt gat tca cac gcc gga aat ctt ccg cgc gaa gga	1584		
163	Gln Val Ala Thr Leu Asp Ser His Ala Gly Asn Leu Pro Arg Glu Gly			
164	515 520 525			
166	gct gta tta att gta acg gcg tct tat aac ggt cat ccg cct gat aac	1632		
167	Ala Val Leu Ile Val Thr Ala Ser Tyr Asn Gly His Pro Pro Asp Asn			
168	530 535 540			
170	gca aag caa ttt gtc gac tgg tta gac caa gcg tct gct gat gaa gta	1680		
171	Ala Lys Gln Phe Val Asp Trp Leu Asp Gln Ala Ser Ala Asp Glu Val			
172	545 550 555			
174	aaa ggc gtt cgc tac tcc gta ttt gga tgc ggc gat aaa aac tgg gct	1728		
175	Lys Gly Val Arg Tyr Ser Val Phe Gly Cys Gly Asp Lys Asn Trp Ala			
176	560 565 570 575			
178	act acg tat caa aaa gtg cct gct ttt atc gat gaa acg ctt gcc gct	1776		
179	Thr Thr Tyr Gln Lys Val Pro Ala Phe Ile Asp Glu Thr Leu Ala Ala			
180	580 585 590			
182	aaa ggg gca gaa aac atc gct gac cgc ggt gaa gca gat gca agc gac	1824		
183	Lys Gly Ala Glu Asn Ile Ala Asp Arg Gly Glu Ala Asp Ala Ser Asp			
184	595 600 605			
186	gac ttt gaa ggc aca tat gaa gaa tgg cgt gaa cat atg tgg agt gac	1872		
187	Asp Phe Glu Gly Thr Tyr Glu Glu Trp Arg Glu His Met Trp Ser Asp			
188	610 615 620			
190	gta gca gcc tac ttt aac ctc gac att gaa aac agt gaa gat aat aaa	1920		
191	Val Ala Ala Tyr Phe Asn Leu Asp Ile Glu Asn Ser Glu Asp Asn Lys			
192	625 630 635			

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/031,146

DATE: 06/05/2002

TIME: 17:09:35

Input Set : A:\031146.txt

Output Set: N:\CRF3\06052002\J031146.raw

```

194 tct act ctt tca ctt caa ttt gtc gac agc gcc gcg gat atg ccg ctt 1968
195 Ser Thr Leu Ser Leu Gln Phe Val Asp Ser Ala Ala Asp Met Pro Leu
196 640 645 650 655
198 gcg aaa atg cac ggt gcg ttt tca acg aac gtc gta gca agc aaa gaa 2016
199 Ala Lys Met His Gly Ala Phe Ser Thr Asn Val Val Ala Ser Lys Glu
200 660 665 670
202 ctt caa cag cca ggc agt gca cga agc acg cga cat ctt gaa att gaa 2064
203 Leu Gln Gln Pro Gly Ser Ala Arg Ser Thr Arg His Leu Glu Ile Glu
204 675 680 685
206 ctt cca aaa gaa gct tct tat caa gaa gga gat cat tta ggt gtt att 2112
207 Leu Pro Lys Glu Ala Ser Tyr Gln Glu Gly Asp His Leu Gly Val Ile
208 690 695 700
210 cct cgc aac tat gaa gga ata gta aac cgt gta aca gca agg ttc ggc 2160
211 Pro Arg Asn Tyr Glu Gly Ile Val Asn Arg Val Thr Ala Arg Phe Gly
212 705 710 715
214 cta gat gca tca cag caa atc cgt ctg gaa gca gaa gaa gaa aaa tta 2208
215 Leu Asp Ala Ser Gln Gln Ile Arg Leu Glu Ala Glu Glu Glu Lys Leu
216 720 725 730 735
218 gct cat ttg cca ctc gct aaa aca gta tcc gta gaa gag ctt ctg caa 2256
219 Ala His Leu Pro Leu Ala Lys Thr Val Ser Val Glu Glu Leu Leu Gln
220 740 745 750
222 tac gtg gag ctt caa gat cct gtt acg cgc acg cag ctt cgc gca atg 2304
223 Tyr Val Glu Leu Gln Asp Pro Val Thr Arg Thr Gln Leu Arg Ala Met
224 755 760 765
226 gct gct aaa acg gtc tgc ccg ccg cat aaa gta gag ctt gaa gcc ttg 2352
227 Ala Ala Lys Thr Val Cys Pro Pro His Lys Val Glu Leu Glu Ala Leu
228 770 775 780
230 ctt gaa aag caa gcc tac aaa gaa caa gtg ctg gca aaa cgt tta aca 2400
231 Leu Glu Lys Gln Ala Tyr Lys Glu Gln Val Leu Ala Lys Arg Leu Thr
232 785 790 795
234 atg ctt gaa ctg ctt gaa aaa tac ccg gcg tgt gaa atg aaa ttc agc 2448
235 Met Leu Glu Leu Leu Glu Lys Tyr Pro Ala Cys Glu Met Lys Phe Ser
236 800 805 810 815
238 gaa ttt atc gcc ctt ctg cca agc ata cgc ccg cgc tat tac tcg att 2496
239 Glu Phe Ile Ala Leu Leu Pro Ser Ile Arg Pro Arg Tyr Tyr Ser Ile
240 820 825 830
242 tct tca tca cct cgt gtc gat gaa aaa caa gca agc atc acg gtc agc 2544
243 Ser Ser Ser Pro Arg Val Asp Glu Lys Gln Ala Ser Ile Thr Val Ser
244 835 840 845
246 gtt gtc tca gga gaa gcg tgg agc gga tat gga gaa tat aaa gga att 2592
247 Val Val Ser Gly Glu Ala Trp Ser Gly Tyr Gly Glu Tyr Lys Gly Ile
248 850 855 860
250 gcg tcg aac tat ctt gcc gag ctg caa gaa gga gat acg att acg tgc 2640
251 Ala Ser Asn Tyr Leu Ala Glu Leu Gln Glu Gly Asp Thr Ile Thr Cys
252 865 870 875
254 ttt att tcc aca ccg cag tca gaa ttt acg ctg cca aaa gac cct gaa 2688
255 Phe Ile Ser Thr Pro Gln Ser Glu Phe Thr Leu Pro Lys Asp Pro Glu
256 880 885 890 895
258 acg ccg ctt atc atg gtc gga ccg gga aca ggc gtc gcg ccg ttt aga 2736

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/031,146

DATE: 06/05/2002

TIME: 17:09:35

Input Set : A:\031146.txt

Output Set: N:\CRF3\06052002\J031146.raw

```

259 Thr Pro Leu Ile Met Val Gly Pro Gly Thr Gly Val Ala Pro Phe Arg
260          900          905          910
262 ggc ttt gtg cag gcg cgc aaa cag cta aaa gaa caa gga cag tca ctt 2784
263 Gly Phe Val Gln Ala Arg Lys Gln Leu Lys Glu Gln Gly Gln Ser Leu
264          915          920          925
266 gga gaa gca cat tta tac ttc ggc tgc cgt tca cct cat gaa gac tat 2832
267 Gly Glu Ala His Leu Tyr Phe Gly Cys Arg Ser Pro His Glu Asp Tyr
268          930          935          940
270 ctg tat caa gaa gag ctt gaa aac gcc caa agc gaa ggc atc att acg 2880
271 Leu Tyr Gln Glu Glu Leu Leu Asn Ala Gln Ser Glu Gly Ile Ile Thr
272          945          950          955
274 ctt cat acc gct ttt tct cgc atg cca aat cag ccg aaa aca tac gtt 2928
275 Leu His Thr Ala Phe Ser Arg Met Pro Asn Gln Pro Lys Thr Tyr Val
276 960          965          970          975
278 cag cac gta atg gaa caa gac ggc aag aaa ttg att gaa ctt ctt gat 2976
279 Gln His Val Met Glu Gln Asp Gly Lys Lys Leu Ile Glu Leu Leu Asp
280          980          985          990
282 caa gga gcg cac ttc tat att tgc gga gac gga agc caa atg gca cct 3024
283 Gln Gly Ala His Phe Tyr Ile Cys Gly Asp Gly Ser Gln Met Ala Pro
284          995          1000          1005
286 gcc gtt gaa gca acg ctt atg aaa agc tat gct gac gtt cac caa gtg 3072
287 Ala Val Glu Ala Thr Leu Met Lys Ser Tyr Ala Asp Val His Gln Val
288          1010          1015          1020
290 agt gaa gca gac gct cgc tta tgg ctg cag cag cta gaa gaa aaa ggc 3120
291 Ser Glu Ala Asp Ala Arg Leu Trp Leu Gln Gln Leu Glu Glu Lys Gly
292          1025          1030          1035
294 cga tac gca aaa gac gtg tgg gct ggg taa 3150
295 Arg Tyr Ala Lys Asp Val Trp Ala Gly
296 1040          1045
298 <210> SEQ ID NO: 2
299 <211> LENGTH: 1048
300 <212> TYPE: PRT
301 <213> ORGANISM: Bacillus megaterium
303 <400> SEQUENCE: 2
305 Thr Ile Lys Glu Met Pro Gln Pro Lys Thr Phe Gly Glu Leu Lys Asn
306 1          5          10          15
308 Leu Pro Leu Leu Asn Thr Asp Lys Pro Val Gln Ala Leu Met Lys Ile
309          20          25          30
311 Ala Asp Glu Leu Gly Glu Ile Phe Lys Phe Glu Ala Pro Gly Arg Val
312          35          40          45
314 Thr Arg Tyr Leu Ser Ser Gln Arg Leu Ile Lys Glu Ala Cys Asp Glu
315          50          55          60
317 Ser Arg Phe Asp Lys Asn Leu Ser Gln Ala Leu Lys Phe Val Arg Asp
318 65          70          75          80
320 Phe Ala Gly Asp Gly Leu Phe Thr Ser Trp Thr His Glu Lys Asn Trp
321          85          90          95
323 Lys Lys Ala His Asn Ile Leu Leu Pro Ser Phe Ser Gln Gln Ala Met
324          100          105          110
326 Lys Gly Tyr His Ala Met Met Val Asp Ile Ala Val Gln Leu Val Gln

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/031,146

DATE: 06/05/2002
TIME: 17:09:36

Input Set : A:\031146.txt
Output Set: N:\CRF3\06052002\J031146.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; N Pos. 16,17,18
Seq#:4; N Pos. 13,14,15
Seq#:5; N Pos. 16,17,18
Seq#:6; N Pos. 17,18,19
Seq#:7; N Pos. 25,26,27
Seq#:8; N Pos. 15,16,17